U.S. Army Research Institute
for the Behavioral and Social Sciences

Research Report 1887

Fidelity Requirements for Army Aviation Training Devices: Issues and Answers

John E. Stewart, David M. Johnson, and William R. Howse
U.S. Army Research Institute

April 2008

U.S. Army Research Institute
for the Behavioral and Social Sciences

A Directorate of the Department of the Army
Deputy Chief of Staff, G1

Authorized and approved for distribution:

Michelle Sams

MICHELLE SAMS
Director

Technical review by

John K. Hawley, Army Research Laboratory,
Human Research and Engineering Directorate, Fort Bliss, TX

Robert T. Nullmeyer, Air Force Research Laboratory, Mesa, AZ

NOTICES

DISTRIBUTION: Primary distribution of this Research Report has been made to ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: DAPE-ARI-ZXM, 2511 Jefferson Davis Highway, Arlington, Virginia 22202-3926

FINAL DISPOSITION: This Research Report may be destroyed with it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings of this Research Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE		
1. REPORT DATE (dd-mm-yy) May 2008	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) May 2007 - January 2008
4. TITLE AND SUBTITLE Fidelity Requirements for Army Aviation Training Devices: Issues and Answers		5a. CONTRACT OR GRANT NUMBER
		5b. PROGRAM ELEMENT NUMBER 622785
6. AUTHOR(S) John E. Stewart, David M. Johnson, and William R. Howse (U.S. Army Research Institute)		5c. PROJECT NUMBER A790
		5d. TASK NUMBER
		5e. WORK UNIT NUMBER 231
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute Fort Rucker Research Unit ATTN: DAPE –ARI-IR Bldg 5100 Fort Rucker, AL 36362-5354		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Institute for the Behavioral & Social Sciences 2511 Jefferson Davis Highway Arlington, VA 22202-3926		10. MONITOR ACRONYM ARI
		11. MONITOR REPORT NUMBER Research Report 1887

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
Subject Matter POC: John E. Stewart II

14. ABSTRACT *(Maximum 200 words)*:

The Future Aviation Simulation Strategies Study Group, sponsored by the U.S. Army Aviation Warfighting Center Directorate of Simulation, presented key questions to the Army Research Institute (ARI) regarding functional requirements (visual, motion, aerodynamic model) for Army helicopter simulators. The present report consists of ARI's responses to these questions, based upon current knowledge of the research. Among the key findings of the report: The prevailing institutional belief is that the simulator, in order to be training effective, must replicate the aircraft. Training consists of offloading flight hours from aircraft to simulator. These assumptions are not supported by scientific evidence. The belief that fidelity equals training effectiveness still drives the acquisition and integration of simulators and training devices. Empirical transfer of training (ToT) experiments using aircraft as criteria are rare. Research has demonstrated that even high fidelity simulators can produce poor ToT to the aircraft, when traditional lock-step training programs are used. Contrariwise, simulators of lesser fidelity have demonstrated acceptable ToT when criterion-based training strategies were employed. The conclusion drawn from the ToT research is that instructional strategies are more important than simulator fidelity. Research on simulator motion shows that while motion may enhance performance in the simulator, it does not seem to impact transfer to the aircraft.

15. SUBJECT TERMS
Flight simulation; helicopter flight training; military aviation training; simulator fidelity; training effectiveness of military aviation simulators; aviation training systems; motion cuing systems; collective training.

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	27	Ellen Kinzer, Technical Publication Specialist 703-602-8047

Standard Form 298

Research Report 1887

Fidelity Requirements for Army Aviation Training Devices: Issues and Answers

John E. Stewart, David M. Johnson, and William R. Howse
U.S. Army Research Institute

Fort Rucker Research Unit
William R. Howse, Chief

U. S. Army Research Institute for the Behavioral and Social Sciences
2511 Jefferson Davis Highway, Arlington, VA 22202-3926

May 2008

Army Project Number
622785A790

Personnel Performance and
Training Technology

ACKNOWLEDGMENTS

The Future Aviation Simulation Strategies (FASS) Study Group (SG) was a farsighted effort convened at Fort Rucker, AL by COL Lee LeBlanc, Director of Simulation and LTC Jerome Peitzman, Chief, Directorate of Simulation Operations and Administration. As a member of the SG, ARI was asked to draw from its own expertise based on decades of aviation training research. The present report is a compilation of ARI's responses to specific questions posed by the SG. ARI's contributions were endorsed by the SG and were incorporated as a major subsection of the FASS Study Report, published 31 July 2007. We at the ARI Fort Rucker Research Unit appreciate this important opportunity to participate in this collaborative effort, which still continues into the next phase of FASS, Spiral Exercises (Spirex). Spirex is dedicated to the development of techniques, tools and procedures for supporting distributed aviation collective training. We would also like to acknowledge the efforts of Alberto Salinas, of Salinas Technologies, who informally and thoroughly reviewed earlier drafts of this report, and provided many helpful comments. Finally, we take our hats off to Dr. John Hawley, of the Army Research Laboratory, Human Research and Engineering Directorate, Fort Bliss, TX, and Dr. Robert Nullmeyer, of the Air Force Research Laboratory, Mesa, AZ, for serving as peer reviewers. They have collaborated with us for years, in the quest to find more efficient and effective means of training Soldiers and Aviators.

FIDELITY REQUIREMENTS FOR ARMY AVIATION TRAINING DEVICES: ISSUES AND ANSWERS

EXECUTIVE SUMMARY

Research Requirement:

As a follow-on to Flight School XXI, the Army's simulation-augmented flight training system, the Directorate of Simulation (DOS) of the U.S. Army Aviation Warfighting Center (USAAWC), organized the Future Aviation Simulation Strategies (FASS) Study Group (SG) to examine the functional requirements for the next generation of simulators and training devices that were to succeed Flight School XXI. The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) at Fort Rucker, AL was asked to provide guidance to the SG, based upon its research experience and extensive knowledge of the research on the fidelity and training requirements of rotary-wing simulators.

Procedure:

The SG presented ARI research staff with a list of key questions regarding simulator fidelity requirements. Many questions, coming from multiple SG participants, were redundant and were combined. Questions were selected which could be answered from existing research data, as opposed to those which would require additional research, or for which answers were not feasible at the present time. Because of the brief (6 month) duration of the FASS SG, ARI's input to the decision process consisted of an extensive review of the open research literature on simulator fidelity and training effectiveness, for both individual and collective training. This procedure culminated in a white paper which attempted to provide thorough, research-based guidance to the SG.

Findings:

The majority of the issues raised by the SG concerned simulator fidelity requirements for effective training, rather than effective instructional strategies. This emphasis on simulation technology revealed a strong belief, prevalent in many aviation training institutions, that the greater the degree of realism created in the virtual environment, the more effective the training. This institutional approach to training consists of offloading flight hours from aircraft to simulator, with training in the simulator being as close as possible to that employed in the aircraft. Every student trains just as in the aircraft, for the same number of hours on a given training day. The assumption that more realism equals better training is not supported by empirical evidence. The research on transfer of training (ToT) from simulator to aircraft has demonstrated that contrary to these institutional beliefs, training strategy has been found to be more important than fidelity with regard to training effectiveness. Lower fidelity simulators have been shown to be training effective when students are trained to proficiency in them; some high fidelity simulators have shown poor ToT when a lock-step, hourly-based training program was employed. Also, research has

shown that when the advantages of the simulator vs. the aircraft are exploited (e.g., successive iterations of a task until it is mastered), training in the simulator plus the aircraft can be superior to training in the aircraft alone. Second to fidelity requirements was the issue as to whether Army helicopter simulators require motion systems. ARI had recently done an extensive review and analysis of this issue, and this information was shared with the SG. The ARI review concluded that although motion can enhance training in the simulator, there is no empirical evidence that the benefits of motion cuing produce ToT to the aircraft. Consequently, the benefits of simulator motion have not been demonstrated, at least with regard to enhancing performance in the aircraft. With regard to virtual training in a collective environment, it was acknowledged that the skills required to be learned were quite different than in the case of individual training, and were better conceptualized as knowledge structures and shared mental models than as psychomotor and rote-learned procedural skills.

Utilization of and Dissemination of Findings:

The conclusions and findings of the present report were incorporated into the FASS Study Report, released 31 July 2007. The findings of the FASS report were briefed to Major General Virgil Packett, Commander, U.S. Army Aviation Warfighting Center, in August, 2007. The Study Report incorporated and endorsed many of the conclusions and recommendations contained in the present ARI report. The most important conclusion, from the standpoint of ARI, is that simulation-focused training strategies have been neglected relative to simulator fidelity requirements, and that more aviation training research (individual and collective) needs to be undertaken in the future. Earlier drafts of this report, in the form of a white paper, were presented to the FASS SG at various stages of the decision process.

FIDELITY REQUIREMENTS FOR ARMY AVIATION TRAINING DEVICES: ISSUES AND ANSWERS

CONTENTS

FIDELITY REQUIREMENTS FOR ARMY AVIATION TRAINING DEVICES: ISSUES AND ANSWERS

Introduction

Background

As a follow-on to Flight School XXI, the Army's simulation-augmented flight training system, the Directorate of Simulation (DOS) of the U.S. Army Aviation Warfighting Center (USAAWC) organized and sponsored a Future Aviation Simulation Strategies (FASS) study group (SG). One critical charge before the FASS SG was to examine the functional requirements for a future family of simulation and training devices that would follow Flight School XXI. Functional requirements include considerations such as motion cuing systems, high-resolution visual display systems, and aerodynamic modeling software. The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) at Fort Rucker, AL was asked to provide input to the SG, based upon its experience in conducting research pertaining to the issues mentioned above.

The questions presented to ARI focused upon issues of fidelity requirements for simulators and implied a relationship to response cueing. The prevailing institutional belief at Fort Rucker, as well as elsewhere in the aviation training community, is that the goal of aviation simulation is to "replicate the aircraft." In addition, there seems to be a belief that training strategies employed in the simulator should be the same as those used in the aircraft—with changes only to accommodate limitations of the synthetic environment. Effective training is assumed because time-honored instructional strategies are employed in a high fidelity virtual environment that closely resembles the aircraft. The problem with these assumptions is that they are not supported by scientific research. However, a veridical synthetic environment that "replicates" the flight characteristics of the target aircraft does not exist. The research literature persuasively indicates that, even if such a virtual world could be implemented, there is no guarantee that it would be training effective. On the other hand, we can be certain that it would be expensive, although future technical developments may reduce costs.

FASS Study Group Milestones

The FASS SG initially convened on November 28, 2006, and met monthly through May 3, 2007. Meetings took place at Fort Rucker, AL. The group's research and fact-finding efforts had a six month duration, culminating in a final report (including fidelity analysis), published July 31, 2007 and decision briefing by Colonel Lee LeBlanc, Director of Simulation, USAAWC, to Major General Virgil Packett, Commanding General, USAAWC and Fort Rucker. The report which follows will concentrate on the specific fidelity and simulator motion issues of which ARI has special knowledge. Other issues before the SG included the technical challenges of interoperating networked

simulators in collective training/ practice environments so that "fair play" between participants is assured, and the evolving standards for semi-automated forces.

Issues and Answers

Fidelity and Training Requirements

The relationship between desired capability and fidelity. One problem with questions of this type is that they assume that benchmarks for simulator fidelity (i.e., cue requirements, simulator complexity) are known. The training developer typically has no objective knowledge of what fidelity is necessary for training specific tasks, and often has to rely on speculation and conjecture. There is not a comprehensive body of scientific research data that specifies the cue requirements for training specific flight tasks, at particular levels (individual, crew, collective), for specific populations of trainees (novice, advanced experienced/recurrent). The task analyses that have been done usually define fidelity requirements based on the subjective and opinions, and their underlying assumptions, elicited from subject matter experts (SMEs) rather than more objective approaches, (e.g., iterations to proficiency in the simulator, transfer of training from the simulator to the aircraft). Furthermore, the SMEs are frequently expert in operation of the target system rather than training development. Thus, we cannot go to any one table or publication and look up a valid answer to this type of question.

Potential negative-habit transfer risk areas associated with a decision not to incorporate a level of required fidelity to overcome a capability gap. The term *negative habit transfer* is a commonly heard component of aviation training jargon. This term confounds the polarity of training effect (diminution of performance instead of improvement) with desirability of the learned behaviors (increases in unwanted behaviors). Both of these are undesired training outcomes but they have different causes and cures. The terms that researchers employ in this sense are positive and negative *transfer of training.* This term relates to whether training experience increases or diminishes performance in the target system, in comparison to a control group which has not received the same training experience. True negative habit transfer is rare. Negative transfer of training is usually found only in laboratory research where the response previously paired with a particular stimulus is drastically changed (i.e., reversed) while the stimulus remains the same. For example, think of what would happen if beginning at midnight tonight, automobile drivers were required to stop at a green light and go at a red one (this was actually ordered by the Red Guard in China during the Cultural Revolution, but they were dissuaded from implementing the resolution by Premier Zhou Enlai). A driver's entire previous driving history would have produced response habits that are incompatible with the present system, and this negative transfer of training could easily be measured by numbers of collisions before and after implementation.

Flight simulators nearly always produce some degree of positive transfer. There is positive transfer even between low fidelity PC-based training devices and the aircraft. This has been well documented in the research literature (e.g., Dennis & Harris, 1998;

Koonce & Bramble, 1998; Talleur, Taylor, Emanuel, Rantanen, & Bradshaw, 2003; Taylor, Lintern, Hulin, Talleur, Emanuel, & Phillips, 1997; Taylor, Talleur, Rantanen, & Emanuel, 2005). Positive transfer of training has been found in the past even for flight simulators such as Link trainers (1-CA-1, GAT-2) and Fort Rucker's own 2B-24 Synthetic Flight Training System that would be considered primitive by today's standards of simulator design (several examples, and references, are presented in the book by Roscoe, 1980). More recent publications continue to report positive transfer of training from simulator to aircraft (Hays, Jacobs, Prince, & Salas, 1992; McCauley, 2006; Patrick, 2003; Stewart & Dohme, 2005; Stewart, Dohme, & Nullmeyer, 2002). The point being made here is: It is not worthwhile worrying about a validated flight simulator produced by a credible vendor causing negative transfer of training.

Common fidelity requirements between individual/crew training and collective training. The problems relating to determining how much simulator complexity is required for training were addressed in detail by Hays et al. (1992), and by Salas, Bowers, and Rhodenizer (1998). Research by Stewart and Dohme (2005), and Stewart et al. (2002), has demonstrated how low-cost simulators can be training effective, when the correct training strategies are employed. One strategy that Stewart and his colleagues found to be effective was the use of proficiency-based training. Student pilots show faster progress when trained individually to an objective performance-based standard for each task than when forced to progress from task to task in a lock-step fashion where all are required to train for a preset number of hours.

The need for simulation-focused training strategies. Hays et al. (1992) state that the way in which simulators are acquired and integrated into training systems, explains why so little progress has been made in determining fidelity requirements. More often than not, the simulators are acquired without knowing their training effectiveness, because no empirical research has been done. The vendors, who manufacture and integrate these devices, do not conduct such research because they are in the business of selling simulators, not research. Occasionally training effectiveness research is conducted after the simulators have been acquired and integrated, but this is narrowly focused on these specific simulators, training specific tasks, in this specific training environment. The research tends to produce no general guidance to the training developer, because of its narrow focus, and because it is conducted on a non-interference basis, making experimental control difficult if not impossible. Hays and his associates also point out that many training developers still believe that the more closely the device resembles the target aircraft, the more training effective it should be. The belief that fidelity equals training effectiveness persists in spite of a body of research evidence showing that, in order to be training effective, a simulator does not necessarily have to resemble the actual operational aircraft (e.g., Wightman & Sistrunk, 1987). In their meta-analytic review of the literature, Hays et al. found only seven helicopter transfer of training (ToT) experiments. There were negligible differences between training in the simulator plus aircraft training vs. aircraft training alone. These authors found a positive ToT effect for fixed-wing aircraft, but not for helicopters, primarily because there were too few helicopter studies that met their criteria for inclusion in the analysis.

Training to proficiency beats lock-step: a case in point. R. T. Nullmeyer (personal communication, January 22, 2007), described his personal experience with an air refueling part task trainer (PTT) for the B-52 (Nullmeyer & Laughery, 1980). This PTT employed a high fidelity CRT display, high-resolution aeromodel, detailed cockpit, and full-motion platform. It was regarded as a high fidelity training device, by early 1980's standards, especially for a PTT. The training syllabus consisted of either five or seven "flights" in the PTT for B-52 pilots taking the aircraft qualification course. Results were not encouraging in either condition. Those students who performed well in the PTT also performed well in "live" air refueling in the B-52. However, those who did not perform well in the PTT did not perform well in the B-52. Consequently, the conclusion was that overall ToT to the aircraft was negligible, in spite of the PTT's state of the art fidelity. Nullmeyer decided after obtaining these preliminary results, to change the training strategy from a fixed number of "flights" for all students, to an individualized strategy based on training to proficiency. The student would train until he reached a criterion of errorless performance to standard (1 continuous 3 min contact). Some reached standard faster than others, but all eventually reached standard and were considered proficient. This resulted in positive ToT to the B-52 (40% fewer sorties to proficiency for copilots upgrading to aircraft commander). The bottom line: even a high fidelity training device can produce poor ToT to the aircraft. This may not have anything to do with simulator fidelity, but with a training program that is not properly tailored to exploit the advantages of the simulator or PTT.

Success story: the U.S. Air Force 58th Special Operations Wing. In a well-designed training program, training in the simulator plus the aircraft should be superior to training in the aircraft alone. This is because the simulator is employed differently than the aircraft, thereby exploiting its training advantages. The Air Force Pave Low program (58th Special Operations Wing) is a success story in its own right (Rakip, Kelly, Appler, & Riley, 1993; Selix, 1993). Selix evaluated the effectiveness of the Pave Low rotary wing training program, which incorporated an integrated suite of training devices, ranging from low fidelity PTTs to full-mission Weapons Systems Trainers (WSTs) for the MH-53J and MH-60L. This investigation was driven by the high operational costs of the MH-53J. In 1986, the MH-53H aircraft qualification course was almost entirely aircraft-based. When the more complex MH-53J replaced the H model in 1990, the course, at 150 training days, became the Air Force's longest aircraft qualification course (AQC), which was unacceptable at a time when flight hours were being reduced.

Thus, it was decided to offload as much training time as possible to simulators and a suite of PTTs and WSTs. Each PTT was dedicated to a specific sensor/ avionics subsystem. Students were trained to proficiency in the least sophisticated device on which the task could be satisfactorily trained. Proficiency had to be demonstrated before the student could proceed to the next level. Once these skills were acquired in the PTTs, they were integrated through crew-level practice in the WST, a full-mission, high fidelity simulator for the MH-53J. Qualification for the Pave Low phase of the course comprised 18 two-hour aircraft sorties. After the introduction of the new training system, this was reduced to 12 sorties in the WST and 3 in the aircraft. Hourly

operational costs for the MH-53J were $3100 vs. approximately $1000 for the WST. Altogether, this resulted in a cost reduction of 70%, for an estimated total savings of $78,300 per student pilot.

The high fidelity MH-53J WST was used as a criterion to validate what was learned in the lower-fidelity training devices. The value added was the hourly costs saved as the training sorties were drastically reduced. Such high fidelity devices could be used for collective training, using DIS/HLA-compliant networking technology. The WST and the OST (Operational System Trainer—a fixed base simulator with high fidelity visuals) were both network-capable. ARI has participated in Digital Training Exercises (DTX) using its OH-58D simulator as a player in the tactical networking scenario. The fixed-base device has high-end PC-based visuals, plus fully functioning instrumentation. It successfully interoperated with a variety of other simulators and PTTs. Questions remain as to the value added by using a higher vs. lower fidelity simulator or trainer. Obviously, part of the payoff is determined by the fidelity requirements of the tasks being practiced.

When the simulator can be better than the aircraft. For some training applications (e.g., reactions to threat) the simulator can have higher fidelity than the aircraft, in that these skills are not practiced in the aircraft in operational training. The WST has proven itself a very effective training environment for these tasks. The following study provides some evidence that a simulation-focused training environment can produce more proficient aircrews than one which depends entirely on the aircraft.

Rakip et al. (1993), conducted a study which validated the claim that, with proficiency based training and selective levels of fidelity, the combination of simulator plus aircraft can yield better results than an aircraft-only training program. When the simulators at Kirtland Air Force Base were down for maintenance, some crews had to revert to training completely in the aircraft. This provided a good opportunity for a "natural experiment" comparing the times it took both groups of crewmembers to qualify when they arrived at their units. Rakip et al. found that new crewmembers trained in the simulator were rated as superior to their aircraft-only counterparts on all criteria except Night Vision Goggle abilities, for which the ratings were virtually the same. Training personnel at the gaining units claimed that simulator-trained crewmembers took two to three months to be brought up to operational standards vs. up to one year for the non simulator-trained crews. Simulator-trained crews required 20 flight hours for combat qualification vs. 50-60 hours for those trained only in the aircraft. The single area where aircraft-only crews performed better was in flying the aircraft, due to more accumulated flight time. Since most of the skills required for the MH-53J were procedural, this tradeoff was considered worthwhile, since procedural skills are forgotten much more rapidly than are perceptual-motor skills. The authors concluded that these post hoc, quasi-experimental results showed that course graduates trained in the PTT/WST coordinated training system exhibited better skill integration and thus progressed faster when they arrived at their units, than those who trained in the MH-53J without simulators and PTTs.

An important point to be gained from these two studies is that the simulator should not be a substitute for the aircraft, but one part of a suite of training tools which can provide superior training, depending on mission, task elements, and instructional strategies. Note that the WST was used for skills integration, after these skills had been acquired on a part task basis in lower fidelity training devices. In this regard, performance in the WST became a criterion for validating the training that had taken place on the PTTs.

Misconceptions about Fidelity and the Use of Simulation

Salas et al. (1998) concluded that the problem with simulation-based training in aviation had little to do with fidelity. Instead, the problem is that while simulation technology has undergone radical change and continues to evolve, training strategies have been frozen in time. They delineated some examples of invalid assumptions in the aviation training community that inhibit effective employment of simulation. These will be paraphrased below.

> *Assumption number one: simulation is all you need.* Though simulation is crucial for learning and practicing flight skills, we must remember that a simulator is simply a tool for training. Training in the simulator should *not* be the same as training in the aircraft. Unfortunately, many training professionals have not gotten this message. One reason for this is that funding emphasizes the simulation technology, but not the learning processes underlying their use. Consequently, the main concerns are the development of synthesized approximations to the real world and not effective, simulation-focused training strategies. The result is usually too much money spent for too little training effectiveness. The authors conclude that the design of training is more important in this context than the fidelity of the simulator.
>
> The quest for a flight simulator that more closely approximates the real world is like a quest for the Holy Grail. Simulation technology, like all digital technology, continues to change and improve every few years. Hence, if simulator technology is driving the acquisition process, all simulators will rapidly become obsolete and will need to be replaced in order to keep up with the state of the art. However, if training effectiveness is the driving concern, then simulators will not become obsolete just because the technology has changed. A flight simulator is simply an environment within which a well-designed program of instruction can be implemented by a competent instructor in order to provide the trainee with an opportunity to learn and practice flight tasks safely.
>
> *Assumption number two: more (fidelity) is better.* Simulation conducted in a high fidelity simulator does not guarantee training success. Some high fidelity simulators have been found not to be training effective, whereas some moderate-to-low fidelity simulators have been found to be effective. Why? The obvious answer is the training program supporting them. More fidelity does not necessarily lead to better learning or to greater ToT to the aircraft. Fidelity by

itself will not produce trained aviators. Simulators, by themselves, do not train. They are tools which, if used as part of a well-designed training program, can allow students to learn and practice aviation tasks.

Assumption number three: if aviators like it, it's good. Salas et al. (1998) believe that this assumption is based upon the ways in which simulators are evaluated—that is by SMEs and trainees. Unfortunately, subjective measures (questionnaires and ratings) are used for these evaluations, which focus on the SME's preconceptions of what a simulator should be and not on the simulation's impact on student learning and performance. Because the simulator is judged more favorably the higher its fidelity, it is simply assumed to be training effective. Immersion in a high fidelity virtual environment, however, does not in itself constitute effective training. Also, compilation of archived performance data is the exception, not the rule. High-fidelity simulators often do not have performance measurement systems built into them. If a researcher wants to collect objective human performance data (based on aircraft state data, not subjective grades or ratings), he or she will have to use a low fidelity simulator. These devices, even PC-based trainers, are able to collect and store performance measures, and they are used frequently in ToT research. Paradoxically, there is a growing body of scientific literature on the training effectiveness of low fidelity simulators, but very little of value on high fidelity simulators. It should be added that experienced aviators seem to be more demanding when it comes to evaluating simulator fidelity and effectiveness (Stewart, 1994, 1995). Stewart found that pilots from an operational unit rated the handling and performance of a high fidelity AH-64 simulator as more similar to the aircraft than did the experienced IPs who participated in the 1995 experiment. Similarly, Stewart, Barker, Weiler, Bonham, and Johnson (2001) found that IPs tended to give lower training effectiveness ratings to a low-cost TH-67 instrument trainer, than did the student pilots who trained in the device. Many of these ratings were based on *perceptions* of fidelity.

The following quote from Salas et al. (1998) concisely sums up what simulation researchers have stated repeatedly:

> We must abandon the notion that simulation equals training and the simplistic view that higher fidelity means better training. As we discussed, these views are not correct and will prevent us from considering and developing more effective strategies for training aviators. (p. 205).

Training trumps fidelity. The principal conclusion of the Salas et al. (1998) analysis, is that fidelity requirements are not as important as the development of well-designed, proficiency-based training programs. A high fidelity simulator with a flawed training program supporting it can be less effective that a low fidelity simulator with a state of the art, proficiency-based training program. We, the authors of this report, concur with this conclusion.

Fidelity and Pilot Experience

Fidelity requirements as a function of pilot experience. This is an important research issue. Recall that we have already stated that IPs seem to be less positive toward the training effectiveness of even high fidelity training devices than the less experienced pilots who train and practice in them. McCauley (2006), in his review of simulator motion cuing requirements, found that experienced pilots not only preferred motion to non motion, but performed better in the simulator (but not the aircraft) when motion was present.

Alessi (2000) and Noble (2002) have taken on the task of addressing the fidelity-experience issue. One important point that they emphasize is the intuitive belief among the training community that more fidelity is universally better than less. The little evidence that exists suggests that this is not true in the case of the novice or student pilot. In fact, at this stage of learning, more fidelity may even be detrimental to effective learning. A high fidelity simulator, or an actual aircraft, presents the student pilot with a multitude of visual, auditory, tactile, and vestibular cues. The student, however, does not need a multitude of sensory cues; he or she needs to learn the critical cues for the tasks being trained. This is not the case for the experienced aviator. High levels of fidelity are appropriate for the expert pilot, who has mastered the critical tasks in his or her assigned aircraft. In this instance new skills are not being acquired; instead skills previously mastered are being validated or refreshed. By contrast, the novice pilot who is still making mistakes and learning through feedback may require a simpler, more focused training environment, more suitable for training these generic, primary skills. At this trainee level of mastery, high fidelity may provide no advantage for ToT to the aircraft, and may introduce too much complexity to an already novel environment. Before determining the level of fidelity, then, one must take into account the level of mastery expected of the learner. Low fidelity training environments are adequate for the novice, whereas experts require (and prefer) higher levels of fidelity. In the latter case, it is not new learning per se but assessment and refreshment of skills which have already been mastered in the aircraft. As familiarization with the simulator and aircraft increases, higher levels of fidelity can be justified. At the present time, there has been much theorizing concerning the simulator fidelity by level of learning relationship, but little actual behavioral research. Why has there been so little research? One reason is the high cost of doing simulator-to-aircraft ToT research.

Simulator Fidelity and the Tasks to be Trained

Reason and methods by which individual crew trainers can support collective training, and effectiveness measures. It will come as no surprise to the aviation training community that the level of simulator fidelity required for training depends on the tasks to be trained. Some examples will make this point clear. Training cockpit procedures does not require a full-mission flight simulator with detailed visuals, cockpit motion, and an aerodynamic software flight model. A cockpit procedures trainer will be sufficient for these tasks. Training Apache pilots in the AQC course how to navigate the AH-64D

multifunction display console, menus, modes, buttons, and knobs does not require a full-mission flight simulator any more than it requires the actual aircraft itself—a PTT dedicated to the multifunction display will satisfy this need. However, what about emergency flight procedures training in AQC or recurrent training at the unit? Emergency flight procedures are by definition dangerous and expensive to reproduce in the actual aircraft. Further, since these tasks are emergency procedures one would want to provide the pilot or crew with a full spectrum of accurate cues in order to train them to diagnose the problem and take appropriate corrective action. This is a case where one can justify a high fidelity, full-mission simulator with accurate flight modeling, motion, warning lights, and flight control feedback.

Fidelity requirements for unit collective training. The level of fidelity required depends upon the tasks to be trained. Crews performing unit collective training can be assumed to know how to fly their aircraft and operate its avionic, navigational, communications, and weapon systems as individual crews. Collective training is for purposes of employing these skills in concert with other aircraft in a company-level or battalion-level exercise that also includes higher headquarters as well as maneuver assets on the ground. In this case would one need the same level of fidelity of the flight systems as in the case where flight itself or emergency procedures were the tasks to be trained? No. But one would presumably want high fidelity communications systems, capable of transmitting both voice and data, as well as a high resolution database with semi-automated forces. Thus, the to-be-trained tasks have a substantial influence upon how much fidelity is provided in which simulator subsystems.

Cost-effective collective training devices. A true collective training system does not have to be as costly or complex as the Aviation Combined Arms Tactical Trainer (AVCATT), a reconfigurable, multi-cockpit collective trainer for Army combat helicopters. By the same token, collective skills do not require a brigade or battalion level combined arms environment in order to be effectively learned and practiced. In the same way that many individual skills can be trained part-task, collective skills can be trained effectively at the platoon or squad level. Training can take place in low-cost, networked simulators or trainers, with a large virtual facility serving as the venue for validation of the training. Units can practice in networked environments at their bases, and, once demonstrating proficiency, can join in simulated maneuvers with other units, and finally, participate in a scored joint exercise at the battalion level. Nor does all of the simulation need to be virtual. Constructive simulation using desktop and laptop devices could serve as a valuable supplement, imparting the cognitive and communication skills required for the development of the shared mental models, which are necessary for successful performance of collective tasks. In short, it may not be necessary to resort to large, expensive training devices in order to model and train the skills required for effective performance in a collective environment.

Cognition, Fidelity, and the Training Environment

Psychomotor vs. cognitive learning. The tasks that are trained in any training environment, whether virtual, constructive, or real, can be divided into two broad

categories: psychomotor (or perceptual-motor) and cognitive. The "stick and rudder" tasks that aviators have had to master ever since the invention of the airplane are psychomotor tasks. They can be learned until they are mastered, then, they can be retained for long periods of time. An analogy is learning to ride a bicycle. We never really forget, once we learn it, but with the passage of time, our skills become degraded. Once we get back on a bike, we are able to refresh these skills quickly. Cognitive tasks are a totally different issue. They require conscious monitoring, are quickly forgotten, and must be refreshed more frequently than psychomotor tasks (Druckman & Bjork, 1991; Goodwin, 2006; Hagman & Rose, 1983; Mengelkoch, Adams, & Gainer, 1971; Sanders, 1999). An example of a cognitive task would be a before-takeoff check or the procedures required to start the engine. One challenge to trainers with the advent of the digital (glass) cockpit technology, is the increase in cognitive monitoring required. This means that frequent practice is necessary, because cognitive skills undergo much more rapid decay than do psychomotor skills. Likewise, the procedures must be intuitive so that pilots will not become "lost" in a complex menu with layers of pages with different functions. Added to the common list of pilot errors is the loss of mode awareness (Lenorovitz, 1990; Sarter & Woods, 1995), in which the crew loses track of the aircraft's flight mode (e.g., takeoff, cruise, descent). This problem came to light in the late 1990s, when two Airbus A-320s, the first commercial aircraft with modern automated cockpits, were lost in crashes attributed to loss of mode awareness.

Not habits, but mental models. With the increasing complexity of aircraft systems and the advent of digital cockpit technology, the majority of tasks that must be learned are cognitive. It is no longer appropriate to refer to "habit" transfer, but "mental models" and "knowledge structures" (Druckman & Bjork, 1991). In virtual simulation environments, procedural tasks are cognitive, and what is being acquired is a mental model of the procedures underlying these tasks. Frequently, these tasks can involve a subsystem of the aircraft, like a sensor or avionics system. As we have already seen from the work of Selix (1993), many of these skills can be trained and sustained part-task, on dedicated PTTs. Desktop/ laptop computers are also candidates for the acquisition and refreshment of cognitively-based tasks.

Simulation Technology in the Collective Training Environment

Networked collective training and mission rehearsal. Much of the training technology in Army Aviation has been concerned with training flight skills to the individual. Once an aviator has mastered these skills, he or she is required to be a functioning member of a crew, which is part of a larger unit, which in turn is part of a much larger organization. One should not lose sight of the fact that aviation missions, like all others, are collective efforts, and that success or failure is measured in the performance of the entire unit. Successful units comprise cohesive combat teams who plan and execute their operations in concert. These concerted actions must be trained and rehearsed in order to maintain currency. This is one facet of training for which emerging virtual training technologies may have great potential (e.g., Nullmeyer & Spiker, 2000).

What is meant by collective training. Collective training in the present context means training a "...composition of aviators involved in multiship operations in which crews operate in separate aircraft but toward the same objectives" (Proctor, Panko, & Donovan, 2004, p. 193). Collective training of Army aviators in the past was accomplished employing multiple actual aircraft, in the so-called "live" flight environment. Obviously, unit level collective training posed a challenge, because it required multiple aircraft, crews, and often, units, as well as a suitable venue for practice. Oftentimes, suitable practice ranges were too distant, or those close by were not suitable for collective practice. Thus proficient aircrews were often frustrated at not being allowed sufficient opportunity for collective unit practice; unit proficiency was not maintained to their satisfaction. Access to opportunities for high-intensity collective practice was limited by accessibility, logistics, time, and cost.

Shared mental models and abstract thinking. Collective training is an area of training research that is still undergoing development. A lot remains to be learned as to just what psychological processes are involved in the learning process, and what methods are best suited for imparting this type of learning. It is in this area that shared mental models (or knowledge structures) seem to play a major role. We are concerned with cognitive pictures of the tactical environment, and the ability of multiple actors to share the same picture of the situation. Likewise, it is important to know how collective decisions are made on the basis of abstract data, when the members of a four-aircraft flight, for example, must decide how to engage an enemy that is not physically present in space and time. By means not yet fully understood, data must become information, and information must become knowledge.

Research on mental models and knowledge structures. The criticality of effective knowledge structures (i.e., mental models) in collective learning situations has been demonstrated by Stout, Salas, and Kraiger (1997). These researchers have pointed out that until recently there has been very little attention paid to shared mental models (complementary knowledge structures) in the aviation training community. Most attempts had used more traditional measures, such as attitude assessments. The Stout et al. experiment attempted, with some success, to determine the effects of team training on the structure of shared knowledge among trainees. Team training was found to result in improved knowledge structures, which, in turn, was found to be a reliable predictor of team performance. The structure of knowledge, to the extent that it is shared by other unit members, should interact with the amount of practice and the amount of information available, to determine performance. It is conceivable that poorly structured knowledge of the situation could lead to worse performance with increasing practice. Theoretical foundations already exist which can provide a conceptual framework for this research. In a sense, we are describing a group decision process, with varying degrees of situational ambiguity, and variations in leadership styles by the mission commander.

Because of the high level of abstraction involved in collective training, it would seem reasonable to suppose that the training and practice of these skills can take place on a variety of networked platforms, ranging from desktop/laptop computers to full-

mission simulators. The simple, low fidelity devices can train the fundamentals of mission planning, where shared mental models must be developed, and where unit members must determine if they actually do share the same models/ knowledge structures about their tactical missions. Learning how to use one's cognitive resources to understand what the team is to do in the future, and projecting one's actions forward in time and space, are activities that impose a heavy demand on the cognitive system, but not necessarily on the hardware and software used to practice these processes.

Finally, team members must learn to make decisions concerning execution of a planned mission, another abstract cognitive process that can be accomplished without a high degree of simulator fidelity. Some promising work on the measurement of situation awareness (SA) by Prince, Ellis, Brannick, and Salas (2007), may provide implications for the development of measures for shared knowledge structures among tactical teams. The researchers address the issue of simulator fidelity for the assessment of implicit and explicit levels of SA. Implicit SA is inferred through observation of overt behavior (Does this aircrew seem to be aware of the situation?), whereas explicit SA is assessed through directing questions to the aircrew (what is the distance and bearing to the target?). Similar to SA, assessing implicit shared knowledge structures may require more complex scenarios and thus more visual and cockpit fidelity than explicit knowledge structures. Nonetheless, practice of scenarios in low fidelity simulators/trainers should transfer to full-mission simulators, and assessments of knowledge structures of tactical teams in the low fidelity scenarios should predict performance in the high fidelity synthetic environment. The tools and techniques for assessing knowledge structures need to be developed and refined; reliable and valid means of assessment are more important than the degree of fidelity of the training system. Prince et al. present some challenging suggestions that could provide guidance in the development of future collective training systems. Some of their conclusions with regard to SA may be applicable to collective training. One recommendation is that low fidelity simulation be used to train student or neophyte pilots in team-level skills. This is because the neophyte's skills have not been developed, and thus require more direct instruction with simpler scenarios. These scenarios are easier to develop and manage in low fidelity simulation environments. For the sustainment and refinement of collective skills, high fidelity environments may be more appropriate, especially for experienced aircrews who are using this practice as a substitute for live company-level exercises. The Prince et al. study suggests that those skills learned in the low fidelity environment will transfer to the high fidelity environment.

Once these processes are mastered, then they could be validated in practice, using different mission scenarios, on higher-fidelity devices. Performance in the latter devices can serve as a means of validating the effectiveness of the training. In brief, the fundamental skills are acquired on low fidelity devices and PTTs, and, once team members have demonstrated proficiency, they and other teams demonstrate what they have learned in a simulation exercise employing networked simulators.

Selective fidelity and functional requirements for collective training. Platt and Crane (1993) pointed out the distinction between functional and physical fidelity. In the

networked trainers used in their research, only those systems required for the collective air combat mission were modeled in high fidelity (e.g., visual display system, aircraft software, throttle and stick). Other systems, such as rudder pedal, various cockpit switches, landing gear and flaps, were absent because they were deemed functionally unnecessary. The importance of this program of research is its relevance to collective training in an aviation-focused environment. It also provides some insight into what is required functionally for particular collective exercises. Lessons learned from the Air Force research could suggest important research issues and approaches for an Army Aviation multi-ship, networked training system.

Yes, but are these systems training effective? The new virtual collective training virtual technology is not just about hardware. AVCATT can create a milieu for multi-ship tactical engagement training and practice, employing a variety of scenarios. It also has considerable potential as a collective training environment, for missions, which units have little opportunity to practice in "live" settings. For example, aircraft types can be "mixed and matched" allowing for missions involving dissimilar aircraft types. One AVCATT trailer also houses a briefing room, which can be used for mission planning and rehearsal as well as for After Action Reviews (AAR). But what are the most efficient and effective strategies for exploiting these assets? At the present time, there are few definitive answers. AVCATT's effectiveness needs to be demonstrated empirically.

Evidence of effectiveness of networked devices. There is some evidence that multi-ship networked training devices have been employed with beneficial results (Bell & Crane, 1993; Crane, Robbins, & Bennett, 2001; Crane, Schiflett, & Oser, 2000; Platt & Crane, 1993). These studies were initiated by the Air Force Research Laboratory (AFRL) Warfighter Training Research Laboratory in Mesa, AZ, which has been investigating the incorporation of Distributed Mission Training (DMT) for four-ship elements, at the level of the operational unit. The AFRL program of research, described in Crane, et al. (2001) sought to determine the functional requirements for such a training system (many of the instructional systems in the older training systems were unwanted, or unused), and to provide unit commanders and instructors with validated instructional strategies and performance measures. The importance of this program of research is its relevance to collective training in an aviation-focused environment. Lessons learned from the Air Force research could suggest important research issues and approaches for an Army Aviation multi-ship, networked training system. Until recently, ARI has also looked at virtual collective training, but primarily in the context of armor and secondarily, infantry.

More recently, a study conducted by the ARI Fort Rucker Research Unit on the current status of helicopter gunnery training within Army Aviation (Sharkey, Stewart, & Salinas, 2005), was initiated through a collaborative effort with DOTD Gunnery Advanced Tactics Branch. This study was an on line survey to which rated Army aviators and non rated crew members who performed airborne gunnery roles were eligible to respond. Many responses to the survey concerned the need for more opportunity for collective training and practice of gunnery skills. Additionally, many

respondents saw a need to bring schoolhouse training more in line with present-day tactical operations. The results of the Sharkey, et al. study prompted DOTD to re-examine current Aircrew Training Manual (ATM) tasks which may no longer reflect the "real world" demands of current combat scenarios in Afghanistan and Iraq. DOTD is concerned that, in addition to not being relevant to present-day tactical demands, some ATM tasks tend to emphasize *individual* skills and proficiencies, when the emphasis should be on crew-level and/or collective training. For example, the activity of participating in a crew-level briefing is conceptualized as an individual activity when it is clearly a crew-level, and, oftentimes, a collective activity.

ARI has undertaken efforts to develop and evaluate networked training of armor forces in a virtual environment (Koger, Long, Britt, Sanders, Broadwater, & Brewer, 1996). The perspective for successful development of any collective training system is to focus on the integration of the system as a whole, and not on its constituent parts. The issues that need to be addressed are not only concerned with traditional transfer of training (virtual to live), but also the capacity for within-simulator (or training environment) skill acquisitions and sustainment. Likewise, research should address whether the skills acquired in the low-cost training devices can transfer to the Combined Arms virtual and/or live environment. In short, do those unit crewmembers that master collective tasks in their networked training environments exhibit superior mission performance in DTX, or in scored live exercises? Also, what alternative instructional strategies seem to hold the most promise for optimizing the effectiveness of this technology?

Matching functional requirements to specific skills to be trained. Although unit commanders have long seen the wisdom behind virtual collective training, questions remain as to the optimal solution to the problem, that is, what would the most efficient and effective collective training system look like? One could envision such as system incorporating the latest visual simulator technology, with state-of-the-art networking, as well as modular, transportable part-task trainers for the practice of procedural tasks, and desktop/laptop constructive simulations to round out and supplement the functional system. But this does not answer certain important, fundamental questions concerning what is needed for the most effective training outcomes at minimal cost. For this reason, and others stated above, it follows that ARI should undertake a comprehensive research program to optimize unit level collective training for Army rotary-wing aviation. This research program will not simply examine functional requirements for collective training systems, but will also delve into theoretical questions concerning what is actually learned in a collective environment (process and mediating variables), as well as the consequences of training (outcome variables). Moreover, it will look closely at what specific skill sets are crucial to successful learning, with the goal of developing performance measures.

Simulator Motion Requirements

Motion requirements for training individual/crew and collective tasks. These issues have been addressed extensively in an ARI Technical Report (McCauley, 2006).

McCauley conducted a review of the research literature, with a focus on whether motion is required for training helicopter pilots. Among the issues addressed by this review were: perceptual fidelity, history of motion bases, disturbance and maneuver motion, human motion sensation, and the empirical evidence for the effectiveness of motion-based flight simulation.

McCauley (2006) found that, while there is a substantial body of empirical evidence to support the effectiveness of flight simulation for training, there is virtually no evidence that supports the training effectiveness of motion platforms. That is, there is no evidence from transfer-of-training experiments conducted in any branch of the aviation community that students trained with simulator motion do any better when transferred to the actual aircraft than students trained in simulators without motion. His review did find, however, that pilots—particularly experienced pilots—perform better *in the simulator* when there is motion present. However, this improvement in in-simulator learning or performance does *not* transfer to the actual aircraft. Likewise, no evidence exists that motion cuing prevents simulator sickness. Unfortunately, a subset of all trainees and instructors will experience simulator sickness until they have adapted to the specific simulator—whether there is motion or not. There is substantial evidence in the research literature supporting the conclusion that pilots prefer motion to no motion when "flying" simulators.

McCauley (2006) concluded that this preference for motion is really a dislike of the static, no-motion-at-all condition. There is evidence that even a quite modest amount of motion, such as provided by a dynamic seat for example, will be acceptable to pilots. Noise, vibration, and other transient cues contribute to the perceived realism of the synthetic experience of flight. McCauley's research efforts also convinced him that the training effectiveness of helicopter flight simulation was more dependent upon thoughtful instructional design than issues of physical fidelity. This conclusion was independently reached by Salas et al. (1998), whose extensive research led them to conclude that fidelity issues were being emphasized and training strategies neglected.

Summary. Simulator motion has not been shown to improve transfer of training to the aircraft. Simulator motion does improve performance while in the simulator—especially for experienced pilots. Motion is not a panacea for simulator sickness. Pilots prefer motion because it seems more "realistic." If training effectiveness is the criterion for incorporating motion into future training simulators, then there are no grounds for doing so. However, training effectiveness, defined as transfer to the aircraft, is only one criterion. Pilots prefer motion, and since pilots are the population that actually uses simulators in their daily work, then the preferences of this population should be considered—pilot acceptance is an important consideration to be weighed in the evaluation and acceptance of aviation training devices.

Tradeoff criteria from a user standpoint, including cost and design tradeoffs. This issue boils down to one of economics. Should the Army invest a substantial portion of a total acquisition budget on a subsystem of unproven effectiveness? Demonstrate one that works first, and then decide whether to buy it. Also, it must be kept in mind that a

motion system does not necessarily require a full six degrees of freedom hexapod. In some instances even a modest motion capability may suffice. ARI's Simulator Training Research Advanced Testbed for Aviation (STRATA) AH-64A simulator employed force cuing in the form of a g-seat. A backward transfer experiment (Stewart, 1994) validated STRATA as having similar handling characteristics to the AH-64A. For this reason, a blanket policy of "no-motion-whatsoever" may not be prudent and may also affect the morale of the pilot population adversely.

Conclusions

One generalization that can be made from the foregoing discussions and analyses is that a paradox exists in the aviation training community; that is, the technological base for simulation continues to evolve at a rapid pace, while the training programs supporting them have shown very little change over the past 20 years. This was evidenced by the fact that ARI-Fort Rucker has dealt with many of the questions contained in the present report over the past decade; very few questions were new or unfamiliar to our research staff. The scientific literature has demonstrated the primacy of proficiency-based training methodology over fidelity, yet the institutional bias in favor of fidelity persists. Part of this may be the perception of the simulator as an attempt to replicate the experience of flight in the aircraft type that it is intended to represent, bolstered by the assumption that the high fidelity simulator can be used as a virtual aircraft. Consequently, flight hours are simply offloaded to the simulator, and training is conducted in the same way as in the aircraft. These assumptions are simply incorrect, and their persistence in the institution of aviation training will not enhance the effectiveness, or more importantly, the *efficiency* of training. Granted, fidelity is important, *depending upon what tasks are to be trained*. However, for most training functions in the simulator, the levels of fidelity currently employed in the full-mission simulators employed in modern training systems such as Flight School XXI, are more than adequate for effective training. The training community, then, should accept the fact that even if modern digital technology could replicate the aircraft, the returns in terms of training outcomes would be disappointing when balanced against the cost of such an investment.

The adoption of new training strategies which are based upon training to proficiency, combined with the integration of high and low fidelity training devices into a total training system, has been shown to yield the best results. Replicating the aircraft (even if it were possible) is unlikely to yield results in which training in simulator plus aircraft is more effective than training in the aircraft alone. *A simulator is not an aircraft, nor should it be used as one.* Considering the capital cost of simulators, it is time to start using them properly as training devices in scientifically-based training programs.

References

Alessi, S. (2000). Simulation design for training and assessment. In H. F. O'Neil, & D. H. Andrews (Eds.), *Aircrew training and assessment* (pp. 197-222). Mahwah, NJ: Erlbaum.

Bell, H. H., & Crane, P. (1993). Training utility of multiship air combat simulation. In G. W. Evans & M. Mollaghasemi (Eds.), *1993 Winter Simulation Conference Proceedings*. Los Angeles: Winter Simulation Conference.

Crane, P., Robbins, R., & Bennett, W. (2001). *Using distributed mission training to augment flight lead upgrade training. (AFRL-HF-AZ-TR-2000-0111*). Mesa, AZ: Air Force Research Laboratory Warfighter Training Research Division.

Crane, P., Schiflett, S.G., & Oser, R.L. (2000). *Roadrunner 98: Training effectiveness in a distributed mission training exercise (AFRL-HE-AZ-TR-2000-0026)*. Mesa, AZ: Air Force Research Laboratory Warfighter Training Research Division.

Dennis, K. A., & Harris, D. (1998). Computer-based simulation as an adjunct to ab initio flight training. *International Journal of Aviation Psychology, 8*(3), 261-276.

Druckman, D., & Bjork, R. A. (Eds.) (1991). *In the mind's eye: Enhancing human performance*. Washington, D.C.: National Academy Press.

Goodwin, G. A. (2006). *The training, retention, and assessment of digital skills: A review and integration of the literature* (ARI Research Report 1864). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Hagman, J. D., & Rose, A. M. (1983). Retention of military tasks: A review. *Human Factors, 25*(2), 199-213.

Hays, R. T., Jacobs, J. W., Prince, C., & Salas, E. (1992). Flight simulator training effectiveness: A meta-analysis. *Military Psychology, 4*, 63-74.

Koger, M.E., Long, D. L., Britt, B. D., Sanders, J. J. Broadwater, T. W., & Brewer, J.D. (1996). *Simulation-based mounted brigade training program: history and lessons learned. Res. Rep. No.* 1689). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Koonce, J. M., & Bramble, W. J. (1998). Personal computer-based flight training devices. *International Journal of Aviation Psychology, 8*(3). 277-292.

Lenorovitz, J. M. (1990, June 25). Indian A320 crash probe data show crew improperly configured aircraft. *Aviation Week & Space Technology, 132*, 84-84.

McCauley, M. E. (2006). *Do Army helicopter training simulators need motion bases?* (ARI Technical Report 1176). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Mengelkoch, R. F., Adams, J. A., & Gainer, C. A. (1971). The forgetting of instrument flying skills. *Human Factors, 13*(5), 397-405.

Noble, C. (2002). The relationship between fidelity and learning in aviation training and assessment. *Journal of Air Transportation, 7*, 33-54.

Nullmeyer, R. T., & Laughery, R. K. (1980). *The effects of ARPTT training on air refueling skill acquisition. Technical Memorandum UDR-TM-80-39*. Dayton, OH: University of Dayton Research Institute.

Nullmeyer, R. T., & Spiker, V. A. (2000). Simulation-based mission rehearsal and human Performance. In H. F. O'Neil, & D. H. Andrews (Eds.). *Aircrew training and assessment* (pp. 131-152). Mahwah, NJ: Erlbaum.

Patrick, J. (2003). Training. In P. S. Tsang, & M. A. Vidulich (Eds.), *Principles and practice of aviation psychology* (pp. 397-434). Mahwah, NJ: Erlbaum.

Platt, P., & Crane, P. (1993). Development, test, and evaluation of a multiship simulation system for air combat training. In *Proceedings of the 15th Industry/ Interservice Training Systems Conference*, Orlando, FL: National Security Industrial Association.

Prince, C., Ellis, E., Brannick, M. T., & Salas. E. (2007). Measurement of team situation awareness in low experience level aviators. *International Journal of Aviation Psychology, 17*, 41-57.

Proctor, M. D., Panko, M., & Donovan, S. J. (2004). Considerations for training team situation awareness and task performance through PC-gamer simulated multiship helicopter operations, *International Journal of Aviation Psychology,* 14, 191-205.

Rakip, R., Kelly, J., Appler, S., & Riley, P. (1993). The role of the MH-53J Weapon System Trainer/ Mission Rehearsal System (WST/MRS) in preparing students for Operation Desert Storm, and future operations. *Proceedings of the 15th Interservice/ Industry Training Systems and Education Conference* (pp. 432-438) Washington, D.C.: American Defense Preparedness Association.

Roscoe, S. N. (1980). *Aviation Psychology.* Ames, Iowa: Iowa State University Press.

Salas, E., Bowers, C. A., & Rhodenizer, L. (1998). It is not how much you have but how you use it: Toward a rational use of simulation to support aviation training. *International Journal of Aviation Psychology, 8*, 197-208

Sanders, W. R. (1999). *Digital procedural skill retention for selected M1A2 tank Inter-Vehicular Information System (IVIS) tasks* (ARI Technical Report 1096). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Sarter, N. B. & Woods, D. D. (1992). Pilot interaction with cockpit automation: Operational experiences with the flight management system. *International Journal of Aviation Psychology, 2*, 303-321.

Selix, G. A. (1993). Evolution of a training program: The effects of simulation on the MH-53J Pave Low Combat Crew Qualification Course. *Proceedings of the 15th Interservice/ Industry Training Systems and Education Conference* (pp. 422-431). Washington, D.C.: American Defense Preparedness Association.

Stewart, J. E. (1994). *Using the backward transfer paradigm to validate the Simulator Training Research Advanced Testbed for Aviation.* (*Res. Rept. No. 1666*). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Stewart, J. E. (1995). *A comparison of two alternative velocity vector cue combinations for the AH-64D Integrated Helmet and Display Sight Subsystem.* (*Res. Rept. No. 1681*). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Stewart, J. E., Barker, W. C., Weiler, D. S., Bonham, J. W., & Johnson, D. M. (2001). *Assessing the effectiveness of a low-cost simulator for instrument training for the TH-67 helicopter* (ARI Research Report 1780). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Stewart, J. E., & Dohme, J. A. (2005). Automated hover trainer: Simulator-based intelligent flight training system. *International Journal of Applied Aviation Studies, 5,* 25-40.

Stewart, J. E., Dohme, J. A., & Nullmeyer, R. T. (2002). U.S. Army initial entry rotary-wing transfer of training research. *International Journal of Aviation Psychology, 12*, 359-375.

Stout, R. J., Salas, E., & Kraiger, K. (1997). The role of trainee knowledge structures in aviation team environments. *International Journal of Aviation Psychology, 7*, 235-250.

Talleur, D. A., Taylor, H. L., Emanuel, T. W., Rantanen, E. M., & Bradshaw, G. L. (2003). Personal computer aviation training devices: Their effectiveness for maintaining instrument currency. *International Journal of Aviation Psychology, 13*, 387-399.

Taylor, H. L., Lintern, G., Hulin, C. L., Talleur, D., Emanuel, T., & Phillips, S. (1997). *Transfer of training effectiveness of personal computer-based aviation training devices* (DOT/FAA/AM-97/11). Washington, DC: Office of Aviation Medicine.

Taylor, H. L., Talleur, D. A., Rantanen, E. M., & Emanuel, T. W. (2005). *The effectiveness of a personal computer aviation training device (PCATD), a flight training device (FTD), and an airplane in conducting instrument proficiency checks*. Paper presented at the 13th International Symposium on Aviation Psychology, Dayton, OH.

Wightman, D. C., & Sistrunk, F. (1987). Part-training strategies in simulated carrier landing final-approach training. *Human Factors, 29,* 245-254.